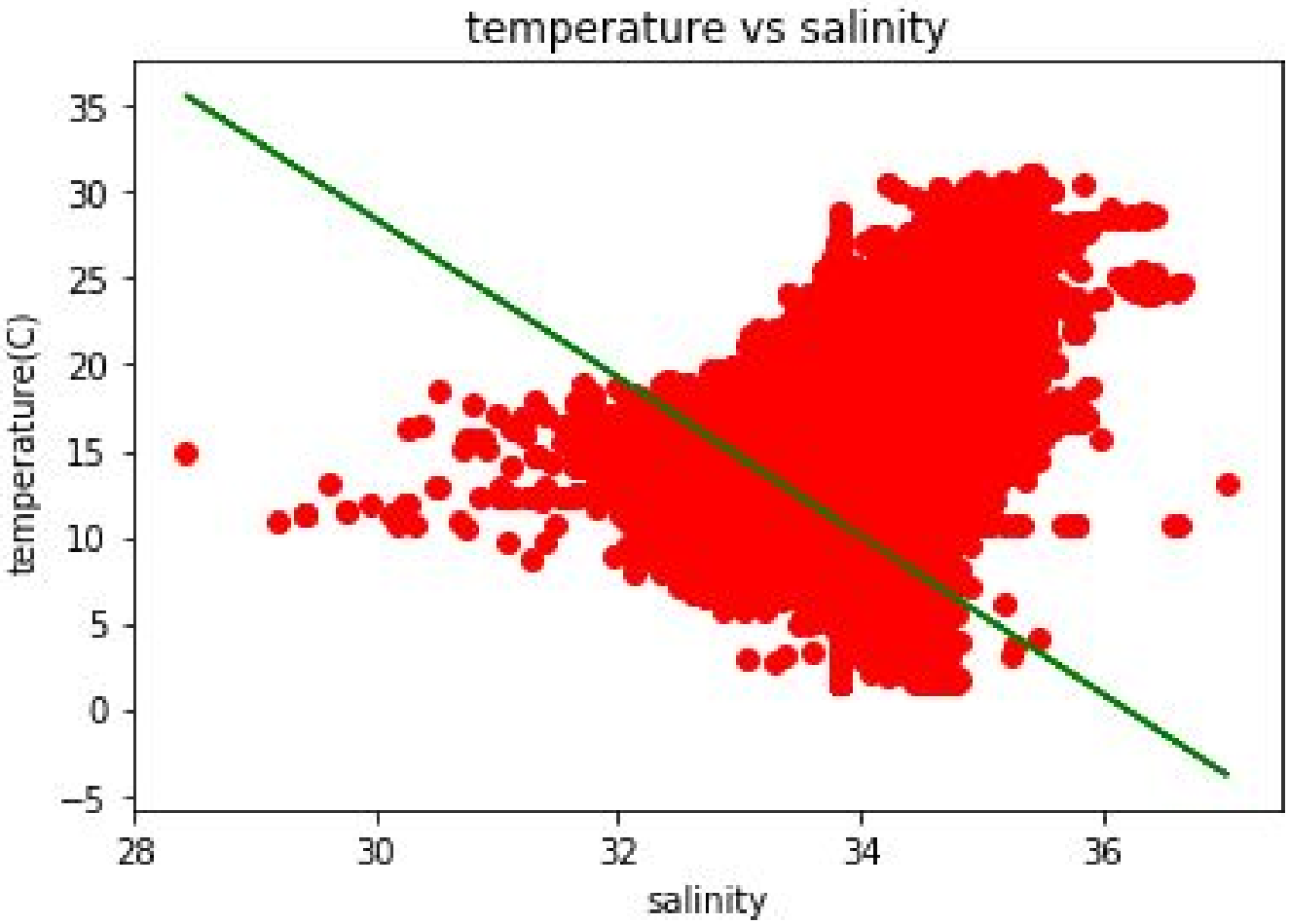
bottle_model.py

Question:
Is there a relationship between water salinity & water temperature? Can you predict the water temperature based on salinity?

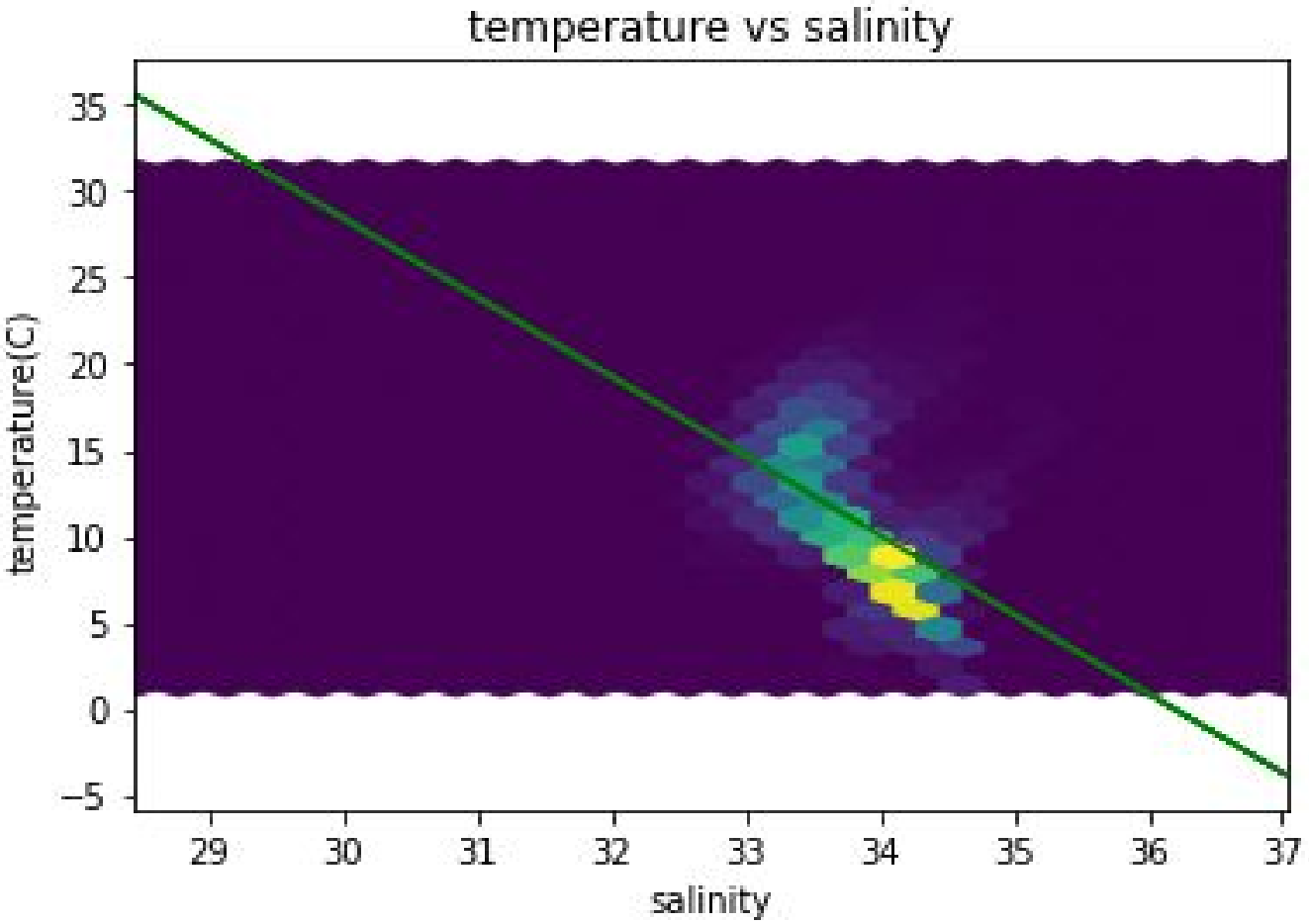
Dataset:
Bottle.csv obtained from https://www.kaggle.com/sohier/calcofi/data

Visual Data:



By looking at graph above , it appears that the linear regression model does not properly represent the data. Since there is a lot of data points we will use a heat map to see if this will help us better data graph obtain

Visual Data(heat map):



In the above heatmap yellow means higher count of point and purple means negligible data point.
By look at our new graph we can now get a better understanding of the relationship between temperature and salinity. With the heatmap we can now see where there is a high concentration of the datapoint.
Note: Since for the missing data we took the average of the other data points we observe a higher concentration of points at these average values

Answer:
We can now observe that as salinity increases we observe an increase in temperature because of the downward relation between the 2 graphs. It appear that we can predict the temperature based on salinity however the accuracy of the model is unknown.